Statistical weaknesses in the alleged RC4 keystream generator

Marina Pudovkina
Moscow Engineering Physics Institute (State University)
maricap@online.ru

Abstract. A large number of stream cipher were proposed and implemented over the last twenty years. In 1987 Rivest designed the RC4 stream cipher, which was based on a different and more software friendly paradigm. It was integrated into Microsoft Windows, Lotus Notes, Apple AOCE, Oracle Secure SQL, and many other applications, and has thus become the most widely used a software-based stream cipher. In this paper we describe some properties of an output sequence of RC4. It is proved that the distribution of first, second output values of RC4 and digraphs are not uniform, which makes RC4 trivial to distinguish between short outputs of RC4 and random strings by analyzing their first, or second output values of RC4 or digraphs.

1. Introduction

A large number of stream cipher were proposed and implemented over the last twenty years. Most of these cipher were based on various combinations of linear feedback shift registers, which were easy to implement in hardware, but relatively slow in software. In 1987 R. Rivest designed the RC4 stream cipher, which was based on a different and more software friendly paradigm. Its design was kept a trade secret until 1994. An anonymous source claimed to have reverse-engineered this algorithm, and published an alleged specification of it in 1994 [1]. It was integrated into Microsoft Windows, Lotus Notes, Apple AOCE, Oracle Secure SQL, and many other applications, and has thus become the most widely used a software-based stream cipher.

The alleged RC4 keystream generator is an algorithm for generating an arbitrarily long pseudorandom sequences based on a variable length key. The pseudorandom sequence is conjectured to be cryptographically secure for using in a stream cipher.

RC4 is in fact a family of algorithms indexed by parameter m, which is a positive integer. The value of m=256 is of greatest interest, as this is value used by all known RC4 applications. In this paper we describe some properties of an output sequence of RC4. It is proved that the distribution of first, second output values of RC4 and digraphs are not uniform. Also we obtain generalizations results of Fluhrer S.R., McGrew D [2] and Mantin I., Shamir A. [3] for different initial values of i_0 and j_0.

The following standard notation will be used throughout:

1. $\boldsymbol{N}$ $=\{1,2,\ldots.\}$,
2. $Z_m=\{0,1,\ldots,m-1\}$,
3. S_m the set of all possible permutations of Z_m.

2. Description of the RC4 cipher

The RC4 stream cipher is modeled a finite automata $A_g= (F, f , Z_m\times Z_m\times S_m, Z_m)$, where F: $Z_m\times Z_m\times S_m \to Z_m\times Z_m\times S_m$ is a next-state function, f: $Z_m\times Z_m\times S_m \to Z_m$ is an output function. The RC4 stream cipher depends on $m=2^n$, $n\in N$.

The state of the RC4 cipher at time t is $(i_t, j_t, s_t)\in Z_m\times Z_m\times S_m$ and the initial state is $(0, 0, s_0,)$. Consider the RC4 cipher at time t ($t=1,2\ldots.$).

The next-state function F

1. $i_t = i_{t-1}+1$ (mod m);
2. $j_t = j_{t-1} + s_{t-1}[i_t]$ (mod m);
3. $s_t[i_t]= s_{t-1}[j_t]$, $s_t[j_t]= s_{t-1}[i_t]$;
4. $s_t[r]= s_{t-1}[r]$, r= $\overline{0, m-1} \backslash \{i_t, j_t\}$.

The output function f
Output: $z_t= s_t[(s_t[j_t]+ s_t[i_t])(\text{mod } m)]$.

Encryption x_t: $c_t=x_t \oplus z_t$. Decryption c_t: $x_t=c_t \oplus z_t$.

3. Description of the used probabilistic model

We will use the following probabilistic model. Assume that the permutation $s_0 \in S_m$ is randomly chosen from S_m, i.e. P{s=s$_0$}=1/m!. Consider a probabilistic model without replacement. Then P{$s_0[r]=a$}=1/m, r=$\overline{0, m-1}$, P{$s_0[r_k]=a_k$ |$s_0[r_{k-1}]=a_{k-1}, \ldots, s_0[r_1]=a_1$}=$\frac{1}{m-k}$, where $\{a_1, \ldots, a_k\} \subseteq \{0, \ldots, m-1\}$, $\{r_1, \ldots, r_k\} \subseteq \{0, \ldots, m-1\}$, $|\{a_1, \ldots, a_k\}|=|\{r_1, \ldots, r_k\}|=k$.

Let us suppose that P{s=s$_1$}=1/m! и P{$s_1[r]=a$}=1/m, $r=\overline{0, m-1}$, P{ $s_1[r_k]=a_k$ | $s_1[r_{k-1}]=a_{k-1}$, …, $s_1[r_1]=a_1$}= $\frac{1}{m-k}$ and s_1 does not depend on *j*.

Note that $j_1= j_0+s_0[i_1]= j_0+s_1[j_1]= j_0+s_0[j_0+s_1[j_1]]$ и γ_1= ($s_0[i_1]+ s_0[j_1]$) (mod m)= ($s_1[i_1]+ s_1[j_1]$) (mod m)= ($s_1[i_1]+ s_1[j_0+s_1[j_1]]$) (mod m).

Proposition 1. *Assume that the permutation $s_0 \in S_m$ is randomly chosen from S_m and γ= ($s_0[i_1]$+ $s_0[j_1]$) (mod m).*
1. If m=1 (mod 2), then

a) P{γ=k}= $\frac{1}{m} - \frac{1}{m(m-1)}$ for k $\neq 2(i_1-j_0)$,

b) P{$\gamma=2(i_1-j_0)$}= $\frac{2}{m}$.

2. If m=0 (mod 2), then

a) P{γ=k}= $\frac{1}{m} - \frac{2}{m(m-1)}$ for k=0 (mod 2), k$\neq 2(i_1-j_0)$,

b) P{γ=k}= $\frac{1}{m}$ for k=1 (mod 2),

c) P{$\gamma=2(i_1–j_0)$}= $\frac{2}{m}$.

Proof. Using γ=($s_0[i_1]+s_0[j_1]$) (mod m)=($s_1[i_1]+ s_1[j_1]$)(mod m)=($s_1[i_1]+s_1[j_0+s_1[j_1]]$)(mod m), we get

$$P\{\gamma=k\}=\sum_{t=0}^{m-1} P\{s_1[i_1]=t\} P\{s_1[j_0+t]=k-t \mid s_1[i_1]=t\} = \sum_{t=0}^{m-1} \frac{1}{m} P\{s_1[j_0+t]=k-t \mid s_1[i_1]=t\} =$$

$$\sum_{\substack{t=0, t \neq i_1-j_0 \\ k \neq 2t}}^{m-1} \frac{1}{m} P\{s_1[j_0+t]=k-t \mid s_1[i_1]=t\} + \frac{1}{m} P\{s_1[i_1]=k-i_1+j_0 \mid s_1[i_1]=i_1-j_0\}.$$

Compute P{γ=k}.

a) If m=1 (mod 2), then for k $\neq 2(i_1-j_0)$ we obtain

$$\mathrm{P}\{\gamma=\mathrm{k}\}= \sum_{\substack{t=0,t\neq i_1-j_0\\ k\neq 2t}}^{m-1} \frac{1}{m} P\{s_1[j_0+t]=k-t \mid s_1[i_1]=t\} = \frac{1}{m}\cdot\frac{m-2}{m-1} = \frac{1}{m} - \frac{1}{m(m-1)}.$$

$$\mathrm{P}\{\gamma=2(\mathrm{i}_1\text{-}\mathrm{j}_0)\}= \sum_{t=0,t\neq i_1-j_0}^{m-1} \frac{1}{m} P\{s_1[j_0+t]=i_1-j_0-t \mid s_1[i_1]=t\} + \frac{1}{m} P\{s_1[i_1]=i_1-j_0 \mid s_1[i_1]=i_1-j_0\} =$$

$$\frac{1}{m}\cdot\frac{m-1}{m-1}+\frac{1}{m}=\frac{2}{m}.$$

Note that $\sum_{k=0}^{m-1}\mathrm{P}\{\gamma=\mathrm{k}\} = \mathrm{P}\{\gamma=\mathrm{k} \mid \mathrm{k} \neq 2(\mathrm{i}_1\text{-}\mathrm{j}_0)\}(\mathrm{m}\text{-}1)+ \mathrm{P}\{\gamma=2(\mathrm{i}_1\text{-}\mathrm{j}_0)\}=1$.

b) If m=0 (mod 2), then

$$\mathrm{P}\{\gamma=2(\mathrm{i}_1\text{-}\mathrm{j}_0)\}= \sum_{t=0,t\neq i_1-j_0}^{m-1} \frac{1}{m} P\{s_1[j_0+t]=i_1-j_0-t \mid s_1[i_1]=t\} + \frac{1}{m} P\{s_1[i_1]=i_1-j_0 \mid s_1[i_1]=i_1-j_0\}$$

$$=\frac{2}{m};$$

$$\mathrm{P}\{\gamma=\mathrm{k} \mid \mathrm{k}=0 \ (\mathrm{mod}\ 2),\ \mathrm{k} \neq 2(\mathrm{i}_1\text{-}\mathrm{j}_0)\}= \sum_{\substack{t=0,t\neq i_1-j_0\\ k\neq 2t}}^{m-1} \frac{1}{m} P\{s_1[j_0+t]=k-t \mid s_1[i_1]=t\} = \frac{1}{m}\cdot\frac{m-3}{m-1} = \frac{1}{m} -$$

$$\frac{2}{m(m-1)};$$

If k=1 (mod 2), then $\mathrm{P}\{\gamma=\mathrm{k}\}= \sum_{t=0,t\neq i_1-j_0}^{m-1} \frac{1}{m} P\{s_1[j_0+t]=k-t \mid s_1[i_1]=t\} = \frac{1}{m}\cdot\frac{m-1}{m-1}=\frac{1}{m}$;

The proof is completed.

4. The distribution of the first output value of RC4

In this section we describe some properties of an output sequence of RC4. It is proved that the distribution of first output values of RC4 is not uniform.

Note that the following result was presented at the MIPT's conference 2001 [9].

Theorem 2. (the distribution of the first output value z_1**)** *Assume that the permutation $s_0 \in S_m$ is randomly chosen from S_m. Let $(i_0, j_0, s_0) \in Z_m \times Z_m \times S_m$ be an initial state of RC4.*

1. *Let $i_1=2j_0$ and for any $m \in \mathbf{N}$:*

a) if $v \neq j_0$, then $P\{z_1=v\}=\frac{1}{m}-\frac{1}{m(m-1)}$,

b) $P\{z_1= j_0\}=\frac{2}{m}$.

2. *Let m=0 (mod 2), i_1=1 (mod 2):*

a) if $v \notin \{ j_0, i_1-j_0\}$, then $P\{z_1=v\} = \frac{1}{m}-\frac{1}{m(m-1)} + \frac{2}{m(m-1)(m-2)}$,

b) $P\{z_1= j_0\}=\frac{2}{m}-\frac{1}{m(m-1)}$,

c) $P\{z_1= i_1-j_0\}=\frac{1}{m}-\frac{2}{m(m-1)}$.

3. *Let m=1 (mod 2):*

a) if $v \notin \{j_0, i_1-j_0, 2^{-1}i_1 (\mathrm{mod}\ m)\}$, then $P\{z_1=v\} = \frac{1}{m}-\frac{1}{m(m-1)} + \frac{3}{m(m-1)(m-2)}$,

b) $P\{z_1= j_0\}=\frac{2}{m}-\frac{1}{m(m-1)}+\frac{1}{m(m-1)(m-2)}$,

c) $P\{ z_1= i_1–j_0\}=\frac{1}{m}-\frac{2}{m(m-1)}+\frac{1}{m(m-1)(m-2)}$,

d) $P\{ z_1=2^{-1}i_1(mod\ m)\}= \frac{1}{m}-\frac{2}{m(m-1)}+\frac{1}{m(m-1)(m-2)}$.

4. *Let* $m=0\ (mod\ 2)$, $i_1=0\ (mod\ 2)$:

a) *if* $\nu\not\in\{j_0,\ i_1–j_0,\ 2^{-1}i_1(mod\ m),\ 2^{-1}i_1+m/2(mod\ m)\}$, *then* $P\{z_1=\nu\}=\frac{1}{m}-\frac{1}{m(m-1)}+\frac{4}{m(m-1)(m-2)}$,

b) $P\{z_1= j_0\}=\frac{2}{m}-\frac{1}{m(m-1)}+\frac{2}{m(m-1)(m-2)}$,

c) $P\{ z_1= i_1–j_0\}=\frac{1}{m}-\frac{2}{m(m-1)}+\frac{2}{m(m-1)(m-2)}$,

d) $P\{ z_1=2^{-1}i_1(mod\ m)\}= \frac{1}{m}-\frac{2}{m(m-1)}+\frac{1}{m(m-1)(m-2)}$,

e) $P\{ z_1=2^{-1}i_1+m/2(mod\ m)\}= \frac{1}{m}-\frac{2}{m(m-1)}+\frac{1}{m(m-1)(m-2)}$.

Proof. Using the full probability formula we obtain that

$$P\{z_1=\nu\}=\sum_{t=0}^{m-1} P\{s_1[i_1]=t\}\sum_{\gamma=0}^{m-1} P\{s_1[\gamma]=\nu \mid s_1[i_1]=t\}\ P\{s_1[t+j_0]=\gamma\text{-}t \mid s_1[i_1]=t,\ s_1[\gamma]=\nu\}= \sum_{\substack{t=0\\ t\neq\nu}}^{m-1} P\{s_1[i_1]=t\}\sum_{\substack{\gamma=0\\ \gamma\neq i_1}}^{m-1} P\{s_1[\gamma]=\nu|s_1[i_1]=t\}\ P\{s_1[t+j_0]=\gamma\text{-}t \mid s_1[i_1]=t, s_1[\gamma]=\nu\}+P\{s_1[i_1]=\nu\}\ P\{s_1[\nu+j_0]= i_1\text{-}\nu \mid s_1[i_1]=\nu\}.$$

Rewrite $P\{z_1=\nu\}=P\{A\}+ P\{B\}$, where

$$P\{A\}=P\{s_1[i_1]=\nu\}\ P\{s_1[\nu+j_0]= i_1\text{-}\nu \mid s_1[i_1]=\nu\} \qquad (1)$$

$$P\{B\}=\sum_{\substack{t=0\\ t\neq\nu}}^{m-1} P\{s_1[i_1]=t\}\sum_{\substack{\gamma=0\\ \gamma\neq i_1}}^{m-1} P\{s_1[\gamma]=\nu \mid s_1[i_1]=t\}\ P\{s_1[t+j_0]= \gamma\text{-}t \mid s_1[i_1]=t, s_1[\gamma]=\nu\} \quad (2)$$

In the following lemma we compute P{A}.

Lemma 1

a) *If* $\nu\neq i_1– j_0$, $i_1\neq 2\nu$, *then* $P\{A\}=\frac{1}{m(m-1)}$,

b) *If either (* $\nu=i_1– j_0$, $i_1\neq 2\nu$ *) or (*$\nu\neq i_1– j_0$, $i_1=2\nu$*), then* $P\{A\}=0$,

c) *If* $\nu = j_0$, $i_1=2j_0$, *then* $P\{A\}=\frac{1}{m}$.

The proof follows from $P\{A\}=P\{s_1[i_1]=\nu\}\ P\{s_1[\nu+j_0]= i_1\text{-}\nu \mid s_1[i_1]=\nu\}$.

Note that P{B} dependences on $P\{s_1[t+j_0]= \gamma\text{-}t \mid s_1[i_1]=t, s_1[\gamma]=\nu\}$, which is

a) if either ($t= i_1–j_0$ и $\gamma=2t$) or ($\gamma–t= j_0$ и $j_0=\nu$), then $P\{s_1[t+j_0]= \gamma\text{-}t \mid s_1[i_1]=t, s_1[\gamma]=\nu\}=P\{ s_1[i_1]=t\}$,

b) if ($t\neq i_1–j_0$ и $\gamma=2t$), ($t=i_1–j_0$ и $\gamma\neq 2t$), ($\gamma–t\neq j_0$ и $\gamma–t=\nu$) and ($\gamma–t= j_0$ и $\gamma–t\neq\nu$), then $P\{s_1[t+j_0]= \gamma–t \mid s_1[i_1]=t, s_1[\gamma]=\nu\}=0$.

Therefore, we can rewrite P{B} as

$$P\{B\}=\sum_{\substack{t=0\\ t\neq\nu}}^{m-1} P\{s_1[i_1]=t\} \sum_{\substack{\gamma=0\\ \gamma\neq i_1}}^{m-1} P\{s_1[\gamma]=\nu \mid s_1[i_1]=t\}\, P\{s_1[t+j_0]=\gamma\text{-}t \mid s_1[i_1]=t,\ s_1[\gamma]=\nu\}=P\{s_1[i_1]= i_1-j_0\}$$

$$P\{s_1[2(i_1-j_0)]=\nu \mid s_1[i_1]= i_1-j_0\}+ P\{ s_1[i_1]= j_0 \mid j_0\neq\nu\} \sum_{\substack{\gamma=0\\ \gamma\neq i_1,2j_0,\\ \gamma\neq j_0+\nu}}^{m-1} P\{s_1[\gamma]=\nu \mid s_1[i_1]= j_0\}\, P\{s_1[2j_0]= \gamma\text{-}\, j_0 \mid$$

$$s_1[i_1]= j_0,\ s_1[\gamma]=\nu\} +P\{s_1[i_1]= i_1-\nu \mid i_1\neq 2\nu,\nu\neq j_0 \} \sum_{\substack{\gamma=0\\ \gamma\neq i_1,2(i_1-\nu),\\ \gamma\neq i_1-\nu+j_0}}^{m-1} P\{s_1[\gamma]=\nu \mid s_1[i_1]= i_1-\nu,\nu\neq j_0\}\, P\{s_1[i_1-$$

$$\nu+j_0]= \gamma\text{-}\, i_1+\nu \mid s_1[i_1]= i_1-\nu,\ s_1[\gamma]=\nu\}+ \sum_{\substack{t=0\\ t\neq\nu,i_1-j_0,\\ t\neq j_0,i_1-\nu}}^{m-1} P\{s_1[i_1]=t\} \sum_{\substack{\gamma=0\\ \gamma\neq i_1,2t,\\ \gamma\neq t+j_0,t+\nu}}^{m-1} P\{s_1[\gamma]=\nu \mid s_1[i_1]=t\}\, P\{s_1[t+j_0]= \gamma\text{-}t \mid$$

$$s_1[i_1]=t,\ s_1[\gamma]=\nu\}+ \sum_{\substack{t=0\\ t\neq(\nu=j_0),i_1-j_0}}^{m-1} P\{s_1[i_1]=t \mid \nu=j_0\}\ P\{s_1[t+j_0]=j_0 \mid s_1[i_1]=t,\ \nu=j_0\}. \quad (3)$$

From (3), we see that P{B} is the sum of four summands, i.e. P{B}= P{B_1}+ P{B_2}+ P{B_3}+ P{B_4}, where

$$P\{B_1\}= P\{s_1[i_1]= i_1-j_0\}\ P\{s_1[2(i_1-j_0)]=\nu \mid s_1[i_1]= i_1-j_0\},$$

$$P\{B_2\}= P\{s_1[i_1]= j_0\} \sum_{\substack{\gamma=0\\ \gamma\neq i_1,2j_0,\\ \gamma\neq j_0+\nu}}^{m-1} P\{s_1[\gamma]=\nu \mid s_1[i_1]= j_0\}\ P\{s_1[2j_0]= \gamma\text{-}\, j_0 \mid s_1[i_1]= j_0,\ s_1[\gamma]=\nu\} \text{ for } j_0\neq\nu,$$

$$P\{B_3\}= P\{s_1[i_1]= i_1-\nu\} \sum_{\substack{\gamma=0\\ \gamma\neq i_1,2(i_1-\nu),\\ \gamma\neq i_1-\nu+j_0}}^{m-1} P\{s_1[\gamma]=\nu \mid s_1[i_1]= i_1-\nu\}\ P\{s_1[i_1-\nu+j_0]= \gamma\text{-}\, i_1+\nu \mid s_1[i_1]= i_1-\nu,$$

$s_1[\gamma]=\nu\}$ for $i_1\neq 2\nu$,

$$P\{B_4\}= \sum_{\substack{t=0\\ t\neq\nu,i_1-j_0,\\ t\neq j_0,i_1-\nu}}^{m-1} P\{s_1[i_1]=t\} \sum_{\substack{\gamma=0\\ \gamma\neq i_1,2t,\\ \gamma\neq t+j_0,t+\nu}}^{m-1} P\{s_1[\gamma]=\nu \mid s_1[i_1]=t\}\ P\{s_1[t+j_0]= \gamma\text{-}t \mid s_1[i_1]=t,\ s_1[\gamma]=\nu\}.$$

P{B_2}=0 for j_0=ν, and P{B_3}=0 for i_1=2ν.

We shall find P{B_1}, P{B_2}, P{B_3}, P{B_4}. Note that P{B_1}, P{B_2}, P{B_3}, P{B_4} accept different values depending on i_0, i_1, j_0, $z_1(\nu)$. For all cases we find P{B_i}, $i=\overline{1,4}$, which are described in the lemmas. For convenience of reading of the proof all these cases are resulted in table 1.

Table 1.
Values P{B_i}, $i=\overline{1,4}$

	$i_1\neq 2j_0$ и $\nu\neq i_1-j_0$	$i_1=2j_0$ or $\nu= i_1-j_0$
P{B_1}	$\frac{1}{m(m-1)}$	0

	($\nu=j_0$), or ($i_1=2j_0$, $\nu\neq j_0$), or ($\nu=i_1$-j_0, $i_1\neq 2j_0$, $j_0\neq\nu$)	$\nu\notin\{i_1$-$j_0, j_0\}$and $i_1\neq 2j_0$
P{B$_2$}	0	$\frac{m-3}{m(m-1)(m-2)}$

	$i_1=2\nu$ or ($\nu=j_0$, $i_1\neq 2\nu$)	$\nu\notin\{i_1–j_0, j_0\}$and $i_1\neq 2\nu$	$\nu= i_1$-j_0, $\nu\neq j_0$, $i_1\neq 2\nu$
P{B$_3$}	0	$\frac{m-3}{m(m-1)(m-2)}$	$\frac{1}{m(m-1)}$

P{B$_4$}	$\nu\notin\{i_1–j_0, j_0\}$and $i_1\neq 2\nu$, $i_1\neq 2j_0$	$\nu=i_1–j_0$, $i_1\neq 2j_0$	$\nu= j_0$, $i_1\neq 2j_0$	$i_1=2\nu$, $\nu\neq j_0$
m=0 (mod 2), i_1=1 (mod 2)	$\frac{(m-4)^2}{m(m-1)(m-2)}$	$\frac{(m-4)}{m(m-1)}$	$\frac{(m-3)}{m(m-1)}$	$\frac{(m-4)(m-3)}{m(m-1)(m-2)}$
m=1 (mod 2)	$\frac{(m-4)^2+1}{m(m-1)(m-2)}$	$\frac{(m-3)^2}{m(m-1)(m-2)}$	$\frac{(m-3)^2}{m(m-1)(m-2)}+\frac{1}{m(m-1)}$	$\frac{(m-4)(m-3)}{m(m-1)(m-2)}$
m=0 (mod 2), i_1=0 (mod 2)	$\frac{(m-4)^2+2}{m(m-1)(m-2)}$	$\frac{(m-3)^2+1}{m(m-1)(m-2)}$	$\frac{(m-3)^2}{m(m-1)(m-2)}+\frac{1}{m(m-1)}+\frac{1}{m(m-1)(m-2)}$	$\frac{(m-4)(m-3)}{m(m-1)(m-2)}$

	$i_1=2j_0$, $\nu\neq j_0$	$\nu= j_0$, $i_1=2j_0$
P{B$_4$}	$\frac{(m-3)^2}{m(m-1)(m-2)}$	$\frac{1}{m}$

Now let us prove the following lemmas.

Lemma 2. *If either $i_1\neq 2j_0$ or $\nu\neq i_1–j_0$, then P{B$_1$}=* $\frac{1}{m(m-1)}$.

Proof. Note that

$$P\{B_1\}= P\{s_1[i_1]= i_1–j_0 \mid \nu\neq i_1–j_0\}\ P\{s_1[2(i_1–j_0)]=\nu \mid s_1[i_1]= i_1–j_0\}=\frac{1}{m(m-1)}.$$

The lemma is proved.

Lemma 3

1. *If $\nu\notin\{i_1–j_0, j_0\}$, then P{B$_2$}=*$\frac{m-3}{m(m-1)(m-2)}$,
2. *If either ($i_1=2j_0$, $\nu\neq j_0$) or ($\nu=i_1$-j_0, $i_1\neq 2j_0$, $j_0\neq\nu$), then P{B$_2$}=0.*

Proof. Let us consider the following cases.

If $\nu\neq i_1$-j_0, $j_0\neq\nu$, then $P\{B_2\}= P\{s_1[i_1]= j_0|\, j_0\neq\nu\} \sum_{\substack{\gamma=0\\ \gamma\neq i_1,2j_0,\\ \gamma\neq j_0+\nu}}^{m-1} P\{s_1[\gamma]=\nu \mid s_1[i_1]= j_0\}\ P\{s_1[2j_0]= \gamma\text{-}\, j_0 \mid s_1[i_1]= j_0,\ s_1[\gamma]=\nu\}=\dfrac{m-3}{m(m-1)(m-2)}$;

If $i_1=2j_0$, $\nu\neq j_0$, then $P\{B_2\}= P\{s_1[2j_0]= j_0|\, j_0\neq\nu\} \sum_{\substack{\gamma=0\\ \gamma\neq 2j_0,\\ \gamma\neq j_0+\nu}}^{m-1} P\{s_1[\gamma]=\nu \mid s_1[2j_0]= j_0\}\ P\{s_1[2j_0]= \gamma\text{-}\, j_0 \mid s_1[2j_0]= j_0,\ s_1[\gamma]=\nu\}=0$.

The lemma is proved.

Lemma 4

1. *If $j_0\neq\nu$, $\nu\neq i_1$-j_0, $i_1\neq 2\nu$, then $P\{B_3\}=\dfrac{m-3}{m(m-1)(m-2)}$,*
2. *If $j_0=\nu$, $i_1\neq 2\nu$, then $P\{B_3\}=0$,*
3. *If $j_0\neq\nu$, $\nu= i_1$-j_0, $i_1\neq 2\nu$, then $P\{B_3\}=\dfrac{1}{m(m-1)}$.*

Proof. Let us consider the following cases. If $j_0\neq\nu$, $\nu\neq i_1$-j_0, $i_1\neq 2\nu$, then

$P\{B_3\}=P\{s_1[i_1]= i_1-\nu|\, i_1\neq 2\nu\} \sum_{\substack{\gamma=0\\ \gamma\neq i_1,2(i_1-\nu),\\ \gamma\neq i_1-\nu+j_0}}^{m-1} P\{s_1[\gamma]=\nu \mid s_1[i_1]= i_1-\nu\}\ P\{s_1[i_1-\nu+j_0]=\gamma-i_1+\nu \mid s_1[i_1]= i_1-\nu,\ s_1[\gamma]= \nu\}= \dfrac{m-3}{m(m-1)(m-2)}$;

If $j_0=\nu$, $i_1\neq 2\nu$, then

$P\{B_3\}=P\{s_1[i_1]= i_1- j_0|\, i_1\neq 2j_0\} \sum_{\substack{\gamma=0\\ \gamma\neq i_1,2(i_1-j_0),}}^{m-1} P\{s_1[\gamma]=j_0 \mid s_1[i_1]= i_1-j_0\}\ P\{s_1[i_1]=\gamma-i_1+j_0 \mid s_1[i_1]= i_1- j_0,\ s_1[\gamma]=j_0\}=0$;

If $j_0\neq\nu$, $\nu= i_1$-j_0, $i_1\neq 2\nu$, then

$P\{B_3\}=P\{s_1[i_1]= j_0|\, i_1\neq 2\nu\} \sum_{\substack{\gamma=0\\ \gamma\neq i_1,2j_0}}^{m-1} P\{s_1[\gamma]=i_1\text{-}j_0 \mid s_1[i_1]= j_0\ \}\ P\{s_1[2j_0]=\gamma- j_0 \mid s_1[i_1]= j_0,\ s_1[\gamma]=i_1\text{-}j_0\}=\dfrac{m-2}{m(m-1)(m-2)}$.

The lemma is proved.

Let us find $P\{B_4\}$. We stress that the value of $P\{B_4\}$ dependents on the number of solutions of $i_1=2t \pmod m$. It follows that we have the following cases, which are dependent on parity of i_1 и m:

a) if $m=0 \pmod 2$, $i_1=1 \pmod 2$, then we have not any solution,
b) if $m=1 \pmod 2$, then we have the following solution $t=2^{-1}i_1 (\mathrm{mod}\ m)$,
c) if $m=0 \pmod 2$, $i_1=0 \pmod 2$, then we have the following solutions $t_1=2^{-1}i_1 (\mathrm{mod}\ m)$ and $t_2=2^{-1}i_1+m/2 \pmod m$.

Therefore, $P\{B_4\}$ and the distribution of the first values dependent on parity of i_1 и m.

Lemma 5. *Let $\nu\neq i_1-j_0$, $\nu\neq j_0$, $i_1\neq 2\nu$, $i_1\neq 2j_0$. Then:*

a) $P\{B_4\}=\dfrac{(m-4)^2}{m(m-1)(m-2)}$ *for m=0 (mod 2), i_1=1 (mod 2),*

b) $P\{B_4\}=\dfrac{(m-4)^2+1}{m(m-1)(m-2)}$ *for m=1 (mod 2),*

c) $P\{B_4\}=\dfrac{(m-4)^2+2}{m(m-1)(m-2)}$ *for m=0 (mod 2), i_1=0 (mod 2), .*

Proof. Let us consider the following cases:

1. if m=0 (mod 2), i_1=1 (mod 2), then

$$P\{B_4\}=\sum_{\substack{t=0\\ t\neq\nu, i_1-j_0,\\ t\neq j_0, i_1-\nu}}^{m-1} P\{s_1[i_1]=t\} \sum_{\substack{\gamma=0\\ \gamma\neq i_1, 2t,\\ \gamma\neq t+j_0, t+\nu}}^{m-1} P\{s_1[\gamma]=\nu \mid s_1[i_1]=t\}\ P\{s_1[t+j_0]=\gamma\text{-}t \mid s_1[i_1]=t,\ s_1[\gamma]=\nu\}=$$

$$\frac{(m-4)^2}{m(m-1)(m-2)}.$$

2. if m=1 (mod 2), then

$$P\{B_4\}=\sum_{\substack{t=0, t\neq 2^{-1}i_1,\\ t\neq\nu, i_1-j_0,\\ t\neq j_0, i_1-\nu}}^{m-1} P\{s_1[i_1]=t\} \sum_{\substack{\gamma=0\\ \gamma\neq i_1, 2t,\\ \gamma\neq t+j_0, t+\nu}}^{m-1} P\{s_1[\gamma]=\nu \mid s_1[i_1]=t\}\ P\{s_1[t+j_0]=\gamma\text{-}t \mid s_1[i_1]=t, s_1[\gamma]=\nu\}+P\{s_1[i_1]=$$

$$2^{-1}i_1\} \sum_{\substack{\gamma=0, \gamma\neq i_1,\\ \gamma\neq 2^{-1}i_1+j_0, 2^{-1}i_1+\nu}}^{m-1} P\{s_1[\gamma]=\nu \mid s_1[i_1]=2^{-1}i_1\}\ P\{s_1[2^{-1}i_1+j_0]=\gamma\text{-}2^{-1}i_1 \mid s_1[i_1]=2^{-1}i_1,$$

$$s_1[\gamma]=\nu\}=\frac{(m-4)(m-5)}{m(m-1)(m-2)}+\frac{(m-3)}{m(m-1)(m-2)}=\frac{(m-4)^2+1}{m(m-1)(m-2)}.$$

3. if m=0 (mod 2), i_1=0 (mod 2), then

$$P\{B_4\}=\sum_{\substack{t=0, t\neq 2^{-1}i_1,\\ t\neq\nu, i_1-j_0, j_0,\\ t\neq i_1-\nu, 2^{-1}i_1+m/2}}^{m-1} P\{s_1[i_1]=t\} \sum_{\substack{\gamma=0\\ \gamma\neq i_1, 2t,\\ \gamma\neq t+j_0, t+\nu}}^{m-1} P\{s_1[\gamma]=\nu \mid s_1[i_1]=t\}\ P\{s_1[t+j_0]=\gamma\text{-}t \mid s_1[i_1]=t,\ s_1[\gamma]=\nu\}$$

$$+P\{s_1[i_1]=2^{-1}i_1\} \sum_{\substack{\gamma=0, \gamma\neq i_1,\\ \gamma\neq 2^{-1}i_1+j_0, 2^{-1}i_1+\nu}}^{m-1} P\{s_1[\gamma]=\nu \mid s_1[i_1]=2^{-1}i_1\}\ P\{s_1[2^{-1}i_1+j_0]=\gamma\text{-}2^{-1}i_1 \mid s_1[i_1]=2^{-1}i_1, s_1[\gamma]=\nu\}$$

$$+P\{s_1[i_1]=2^{-1}i_1+m/2\} \sum_{\substack{\gamma=0, \gamma\neq i_1,\\ \gamma\neq 2^{-1}i_1+j_0+m/2,\\ \gamma\neq 2^{-1}i_1+\nu+m/2}}^{m-1} P\{s_1[\gamma]=\nu \mid s_1[i_1]=2^{-1}i_1+m/2\}\ P\{s_1[2^{-1}i_1+m/2+j_0]=\gamma\text{–}m/2+2^{-1}i_1 \mid$$

$$s_1[i_1]=2^{-1}i_1+m/2, s_1[\gamma]=\nu\}=\frac{(m-4)(m-6)}{m(m-1)(m-2)}+\frac{2(m-3)}{m(m-1)(m-2)}=\frac{(m-4)^2+2}{m(m-1)(m-2)}.$$

The lemma is proved.

Lemma 6. *Let $\nu=i_1–j_0$, $i_1\neq 2j_0$. Then:*

a) $P\{B_4\}=\dfrac{(m-4)}{m(m-1)}$ *for m=0 (mod 2), i_1=1 (mod 2),*

b) $P\{B_4\}=\dfrac{(m-3)^2}{m(m-1)(m-2)}$ *for m=1 (mod 2),*

c) $P\{B_4\}=\dfrac{(m-3)^2+1}{m(m-1)(m-2)}$ *for m=0 (mod 2), i_1=0 (mod 2).*

Proof. Note that

$$P\{B_4\}= \sum_{\substack{t=0\\ t\neq i_1-j_0,\, j_0}}^{m-1} P\{s_1[i_1]=t\} \sum_{\substack{\gamma=0\\ \gamma\neq i_1, 2t,\\ \gamma\neq t+j_0, t+i_1-j_0}}^{m-1} P\{s_1[\gamma]=i_1-j_0 \mid s_1[i_1]=t\}\, P\{s_1[t+j_0]= \gamma\text{-}t \mid s_1[i_1]=t,\ s_1[\gamma]=i_1-j_0\}$$

and consider the following cases:

1. if m=0 (mod 2), i_1=1 (mod 2), then

$$P\{B_4\}=\frac{(m-4)(m-2)}{m(m-1)(m-2)}=\frac{(m-4)}{m(m-1)}.$$

2. if m=1 (mod 2), then

$$P\{B_4\}= \sum_{\substack{t=0, t\neq 2^{-1}i_1,\\ t\neq i_1-j_0,\, j_0}}^{m-1} P\{s_1[i_1]=t\} \sum_{\substack{\gamma=0\\ \gamma\neq i_1, 2t,\\ \gamma\neq t+j_0, t+\nu}}^{m-1} P\{s_1[\gamma]=i_1-j_0 \mid s_1[i_1]=t\}\ P\{s_1[t+j_0]= \gamma\text{-}t \mid s_1[i_1]=t,\ s_1[\gamma]=i_1-$$

$$j_0\}+P\{s_1[i_1]= 2^{-1}i_1\} \sum_{\substack{\gamma=0, \gamma\neq i_1,\\ \gamma\neq 2^{-1}i_1+j_0, 2^{-1}i_1+\nu}}^{m-1} P\{s_1[\gamma]=i_1-j_0 \mid s_1[i_1]=2^{-1}i_1\}\ P\{s_1[2^{-1}i_1+j_0]= \gamma\text{-}2^{-1}i_1 \mid s_1[i_1]= 2^{-1}i_1,$$

$$s_1[\gamma]=i_1-j_0\}=\frac{(m-4)(m-3)}{m(m-1)(m-2)}+\frac{(m-3)}{m(m-1)(m-2)}=\frac{(m-3)^2}{m(m-1)(m-2)};$$

3. if m=0 (mod 2), i_1=0 (mod 2), then

$$P\{B_4\}= \sum_{\substack{t=0, t\neq 2^{-1}i_1,\\ t\neq i_1-j_0,\, j_0,\\ t\neq 2^{-1}i_1+m/2}}^{m-1} P\{s_1[i_1]=t\} \sum_{\substack{\gamma=0\\ \gamma\neq i_1, 2t,\\ \gamma\neq t+j_0, t+\nu}}^{m-1} P\{s_1[\gamma]=\nu \mid s_1[i_1]=t\}\ P\{s_1[t+j_0]=\gamma\text{-}t \mid s_1[i_1]=t,\ s_1[\gamma]=\nu\}+ P\{s_1[i_1]=$$

$$2^{-1}i_1\} \sum_{\substack{\gamma=0, \gamma\neq i_1,\\ \gamma\neq 2^{-1}i_1+j_0, 2^{-1}i_1+\nu}}^{m-1} P\{s_1[\gamma]=\nu \mid s_1[i_1]=2^{-1}i_1\}\ P\{s_1[2^{-1}i_1+j_0]=\gamma\text{-}2^{-1}i_1 \mid s_1[i_1]=2^{-1}i_1,\ s_1[\gamma]=\nu\}+ P\{s_1[i_1]= 2^{-1}$$

$$i_1+m/2\} \sum_{\substack{\gamma=0, \gamma\neq i_1,\\ \gamma\neq 2^{-1}i_1+j_0+m/2,\\ \gamma\neq 2^{-1}i_1+\nu+m/2}}^{m-1} P\{s_1[\gamma]=\nu \mid s_1[i_1]=2^{-1}i_1+m/2\}\ P\{s_1[2^{-1}i_1+m/2+j_0]= \gamma-m/2+2^{-1}i_1 \mid s_1[i_1]= 2^{-1}i_1+m/2,$$

$$s_1[\gamma]=\nu\}=\frac{(m-4)^2}{m(m-1)(m-2)}+\frac{2(m-3)}{m(m-1)(m-2)}=\frac{(m-3)^2+1}{m(m-1)(m-2)}.$$

The lemma is proved.

Lemma 7. *Let $\nu= j_0$, $i_1\neq 2j_0$. Then:*

a) $P\{B_4\}=\dfrac{(m-3)}{m(m-1)}$ *for m=0 (mod 2), i_1=1 (mod 2),*

b) $P\{B_4\}=\dfrac{(m-3)^2}{m(m-1)(m-2)}+\dfrac{1}{m(m-1)}$ *for m=1 (mod 2),*

c) $P\{B_4\}=\dfrac{(m-3)^2}{m(m-1)(m-2)}+\dfrac{1}{m(m-1)}+\dfrac{1}{m(m-1)(m-2)}$ *for m=0 (mod 2), i_1=0 (mod 2).*

Proof. Note that

$$P\{B_4\}= \sum_{\substack{t=0\\ t\neq i_1-j_0,\, j_0,}}^{m-1} P\{s_1[i_1]=t\} \sum_{\substack{\gamma=0\\ \gamma\neq i_1, 2t,\\ \gamma\neq t+j_0}}^{m-1} P\{s_1[\gamma]=j_0 \mid s_1[i_1]=t\}\ P\{s_1[t+j_0]= \gamma\text{-}t \mid s_1[i_1]=t,\ s_1[\gamma]=j_0\} \text{ and}$$

consider the following cases:

1. if m=0 (mod 2), i_1=1 (mod 2), then

$$P\{B_4\}=\frac{(m-3)(m-2)}{m(m-1)(m-2)}=\frac{(m-3)}{m(m-1)}.$$

2. if m=1 (mod 2), то

$$P\{B_4\}=\sum_{\substack{t=0,t\neq 2^{-1}i_1,\\ t\neq i_1-j_0,j_0}}^{m-1} P\{s_1[i_1]=t\}\sum_{\substack{\gamma=0\\ \gamma\neq i_1,2t,\\ \gamma\neq t+j_0,}}^{m-1} P\{s_1[\gamma]=j_0 \mid s_1[i_1]=t\}\; P\{s_1[t+j_0]=\gamma\text{-}t \mid s_1[i_1]=t,\ s_1[\gamma]=j_0\}+P\{s_1[i_1]=2^{-1}$$

$$i_1\}\sum_{\substack{\gamma=0,\gamma\neq i_1,\\ \gamma\neq 2^{-1}i_1+j_0,}}^{m-1} P\{s_1[\gamma]=j_0 \mid s_1[i_1]=2^{-1}i_1\}\; P\{s_1[2^{-1}i_1+j_0]=\gamma\text{-}2^{-1}i_1 \mid s_1[i_1]=2^{-1}i_1,\ s_1[\gamma]=j_0\}=$$

$$\frac{(m-3)^2}{m(m-1)(m-2)}+\frac{(m-2)}{m(m-1)(m-2)}=\frac{(m-3)^2}{m(m-1)(m-2)}+\frac{1}{m(m-1)};$$

3. if m=0 (mod 2), i_1=0 (mod 2), then

$$P\{B_4\}=\sum_{\substack{t=0,t\neq 2^{-1}i_1,\\ t\neq i_1-j_0,j_0,\\ t\neq 2^{-1}i_1+m/2}}^{m-1} P\{s_1[i_1]=t\}\sum_{\substack{\gamma=0\\ \gamma\neq i_1,2t,\\ \gamma\neq t+j_0}}^{m-1} P\{s_1[\gamma]=j_0 \mid s_1[i_1]=t\}\; P\{s_1[t+j_0]=\gamma\text{-}t \mid s_1[i_1]=t,\ s_1[\gamma]=j_0\}+P\{s_1[i_1]=$$

$$2^{-1}i_1\}\sum_{\substack{\gamma=0,\gamma\neq i_1,\\ \gamma\neq 2^{-1}i_1+j_0,}}^{m-1} P\{s_1[\gamma]=j_0 \mid s_1[i_1]=2^{-1}i_1\}\; P\{s_1[2^{-1}i_1+j_0]=\gamma\text{-}2^{-1}i_1 \mid s_1[i_1]=2^{-1}i_1,\ s_1[\ \gamma]=j_0\}+P\{s_1[i_1]=2^{-1}\ i_1+$$

$$m/2\}\sum_{\substack{\gamma=0,\gamma\neq i_1,\\ \gamma\neq 2^{-1}i_1+j_0+m/2,}}^{m-1} P\{s_1[\gamma]=j_0 \mid s_1[i_1]=2^{-1}i_1+m/2\}\; P\{s_1[2^{-1}i_1+m/2+j_0]=\gamma\text{–}m/2+2^{-1}i_1 \mid s_1[i_1]=2^{-1}i_1+m/2,$$

$$s_1[\gamma]=j_0\}=\frac{(m-4)(m-3)}{m(m-1)(m-2)}+\frac{2(m-2)}{m(m-1)(m-2)}=\frac{(m-3)^2}{m(m-1)(m-2)}+\frac{1}{m(m-1)}+\frac{1}{m(m-1)(m-2)}.$$

The lemma is proved.

Lemma 8. *Let $i_1=2\nu$, $\nu\neq j_0$. Then $P\{B_4\}=$* $\frac{(m-4)(m-3)}{m(m-1)(m-2)}$.

Proof. For m=1(mod 2) we have

$$P\{B_4\}=\sum_{\substack{t=0,t\neq j_0\\ t\neq \nu,2\nu-j_0,}}^{m-1} P\{s_1[i_1]=t\}\sum_{\substack{\gamma=0\\ \gamma\neq 2\nu,2t,\\ \gamma\neq t+j_0,t+\nu}}^{m-1} P\{s_1[\gamma]=\nu \mid s_1[i_1]=t\}\; P\{s_1[t+j_0]=\gamma\text{-}t \mid s_1[i_1]=t,\ s_1[\gamma]=\nu\}=$$

$$\frac{(m-4)(m-3)}{m(m-1)(m-2)}=\frac{1}{m}-\frac{5}{m(m-1)}+\frac{2}{m(m-1)(m-2)}.$$

As before we prove another cases. The lemma is proved.

Lemma 9 *Let $i_1=2j_0$. Then:*

a) $P\{B_4\}=\frac{(m-3)^2}{m(m-1)(m-2)}$ for $\nu\neq j_0$,

b) $P\{B_4\}=\frac{1}{m}$ for $\nu= j_0$.

Proof. Let us consider the following cases:

1. If $\nu\neq j_0$, then

P{B$_4$}= $\sum_{\substack{t=0,t\neq\nu\\ t\neq j_0,2j_0-\nu}}^{m-1}$ P{s$_1$[i$_1$]=t} $\sum_{\substack{\gamma=0\\ \gamma\neq 2j_0,2t,\\ \gamma\neq t+j_0}}^{m-1}$ P{s$_1$[γ]=ν | s$_1$[i$_1$]=t} P{s$_1$[t+j$_0$]=γ–t | s$_1$[i$_1$]=t, s$_1$[γ]=ν}=

$$\frac{(m-3)^2}{m(m-1)(m-2)}.$$

2. if ν= j$_0$, then

P{B$_4$}= $\sum_{t=0,t\neq j_0}^{m-1}$ P{s$_1$[2j$_0$]=t} $\sum_{\substack{\gamma=0\\ \gamma\neq 2j_0,\\ \gamma\neq t+j_0}}^{m-1}$ P{s$_1$[γ]=j$_0$ | s$_1$[2j$_0$]=t} P{s$_1$[t+j$_0$]=γ–t | s$_1$[2j$_0$]=t, s$_1$[γ]=j$_0$}=

$$\frac{(m-1)(m-2)}{m(m-1)(m-2)}=\frac{1}{m}.$$

The lemma is proved.

Now we compute the distribution of the first output value of RC4, i.e. P{z$_1$=ν}. This will be complete the proof.

Lemma 10. *Let ν≠i$_1$–j$_0$, ν≠ j$_0$, i$_1$≠2ν, i$_1$≠2j$_0$. Then:*

a) P{ z$_1$=ν}= $\frac{1}{m}-\frac{1}{m(m-1)}+\frac{2}{m(m-1)(m-2)}$ *for m=0 (mod 2), i$_1$=1 (mod 2),*

b) P{z$_1$=ν}= $\frac{1}{m}-\frac{1}{m(m-1)}+\frac{3}{m(m-1)(m-2)}$ *for m=1 (mod 2),*

c) P{z$_1$=ν}= $\frac{1}{m}-\frac{1}{m(m-1)}+\frac{4}{m(m-1)(m-2)}$ *for m=0 (mod 2), i$_1$=0 (mod 2).*

Proof. Note that

P{z$_1$=ν}=P{A}+ P{B$_1$}+ P{B$_2$}+P{B$_3$}+P{B$_4$}= $\frac{1}{m(m-1)}+\frac{1}{m(m-1)}+\frac{m-3}{m(m-1)(m-2)}+\frac{m-3}{m(m-1)(m-2)}$ + P{B$_4$}= $\frac{4}{m(m-1)}-\frac{2}{m(m-1)(m-2)}$ + P{B$_4$} and consider the following cases:

1. if m=0 (mod 2), i$_1$=1 (mod 2), then

P{z$_1$=ν}= $\frac{4}{m(m-1)}-\frac{2}{m(m-1)(m-2)}+\frac{(m-4)^2}{m(m-1)(m-2)}=\frac{m^2-4m+6}{m(m-1)(m-2)}=\frac{1}{m}-\frac{1}{m(m-1)}+\frac{2}{m(m-1)(m-2)}.$

2. if m=1 (mod 2), then

P{z$_1$=ν}= $\frac{4}{m(m-1)}-\frac{2}{m(m-1)(m-2)}+\frac{(m-4)^2+1}{m(m-1)(m-2)}=\frac{1}{m}-\frac{1}{m(m-1)}+\frac{3}{m(m-1)(m-2)}.$

3. if m=0 (mod 2), i$_1$=0 (mod 2), then

P{z$_1$=ν}= $\frac{4}{m(m-1)}-\frac{2}{m(m-1)(m-2)}+\frac{(m-4)^2+2}{m(m-1)(m-2)}=\frac{1}{m}-\frac{1}{m(m-1)}+\frac{4}{m(m-1)(m-2)}.$

The lemma is proved.

Lemma 11. *Let* $\nu=i_1–j_0$, $i_1\neq 2j_0$. *Then:*

a) $P\{ z_1= i_1–j_0\}=\frac{1}{m}-\frac{2}{m(m-1)}$ *for m=0 (mod 2),* $i_1=1$ *(mod 2),*

b) $P\{ z_1= i_1–j_0\}= \frac{1}{m}-\frac{2}{m(m-1)}+\frac{1}{m(m-1)(m-2)}$ *for m=1 (mod 2),*

c) $P\{z_1= i_1–j_0\}=\frac{1}{m}-\frac{2}{m(m-1)}+\frac{2}{m(m-1)(m-2)}$ *for m=0 (mod 2),* $i_1=0$ *(mod 2).*

Proof. Note that

$P\{z_1= i_1–j_0\}= P\{A\}+ P\{B_1\}+ P\{B_2\}+P\{B_3\}+P\{B_4\}=0 +0+\frac{1}{m(m-1)}+ P\{B_4\}$ and consider the following cases:

1. if m=0 (mod 2), $i_1=1$ (mod 2), then

$$P\{z_1= i_1–j_0\}=\frac{1}{m(m-1)}+\frac{(m-4)}{m(m-1)}=\frac{1}{m}-\frac{2}{m(m-1)}.$$

2. if m=1 (mod 2), then

$$P\{z_1= i_1–j_0\}=\frac{1}{m(m-1)}+\frac{(m-3)^2}{m(m-1)(m-2)}=\frac{1}{m}-\frac{2}{m(m-1)}+\frac{1}{m(m-1)(m-2)}.$$

3. if m=0 (mod 2), $i_1=0$ (mod 2), then

$$P\{z_1= i_1–j_0\}=\frac{1}{m(m-1)}+\frac{(m-3)^2+1}{m(m-1)(m-2)}=\frac{1}{m}-\frac{2}{m(m-1)}+\frac{2}{m(m-1)(m-2)}.$$

The lemma is proved.

Lemma 12. *Let* $\nu= j_0$, $i_1\neq 2j_0$. *Then:*

a) $P\{z_1= j_0\}=\frac{2}{m}-\frac{1}{m(m-1)}$ *for m=0 (mod 2),* $i_1=1$ *(mod 2),*

b) $P\{ z_1= j_0\}=\frac{2}{m}-\frac{1}{m(m-1)}+\frac{1}{m(m-1)(m-2)}$ *for m=1 (mod 2),*

c) $P\{z_1= j_0\}=\frac{2}{m}-\frac{1}{m(m-1)}+\frac{2}{m(m-1)(m-2)}$ *for m=0 (mod 2),* $i_1=0$ *(mod 2).*

Proof. Note that

$$P\{z_1= j_0\}= \sum_{\substack{t=0\\ t\neq(\nu=j_0),i_1-j_0}}^{m-1} P\{s_1[i_1]=t \mid \nu=j_0\}\ P\{s_1[t+j_0]=j_0 \mid s_1[i_1]=t, \nu=j_0\}+P\{A\}+ P\{B_1\}+ P\{B_2\}+P\{B_3\}$$

$$+P\{B_4\}=\frac{(m-2)}{m(m-1)}+\frac{1}{m(m-1)}+\frac{1}{m(m-1)}+0+0+P\{B_4\}= \frac{2}{m(m-1)}+\frac{(m-2)}{m(m-1)}+P\{B_4\}.$$

Let us consider the following cases:

1. if m=0 (mod 2), $i_1=1$ (mod 2), then

$$P\{z_1= j_0\}=\frac{2}{m(m-1)}+\frac{(m-3)}{m(m-1)}+\frac{(m-2)}{m(m-1)}=\frac{2}{m}-\frac{1}{m(m-1)}.$$

2. if m=1 (mod 2), then

$$P\{z_1= j_0\}=\frac{2}{m(m-1)}+\frac{(m-2)}{m(m-1)}+\frac{(m-3)^2}{m(m-1)(m-2)}+\frac{1}{m(m-1)}=\frac{1}{m}+\frac{1}{m(m-1)}+\frac{1}{m}-\frac{2}{m(m-1)}+$$
$$\frac{1}{m(m-1)(m-2)}=\frac{2}{m}-\frac{1}{m(m-1)}+\frac{1}{m(m-1)(m-2)}.$$

3. if m=0 (mod 2), $i_1=0$ (mod 2), then

$$P\{z_1=j_0\}=\frac{(m-2)}{m(m-1)}+\frac{2}{m(m-1)}+\frac{(m-3)^2}{m(m-1)(m-2)}+\frac{1}{m(m-1)}+\frac{1}{m(m-1)(m-2)}=\frac{2}{m}-\frac{1}{m(m-1)}+\frac{2}{m(m-1)(m-2)}.$$

The lemma is proved.

Lemma 13. *Let $i_1=2v$, $v\neq j_0$. Then*

$$P\{z_1=v\}=\frac{1}{m}-\frac{2}{m(m-1)}+\frac{1}{m(m-1)(m-2)}.$$

Proof. Note that

$$P\{z_1=v\}=P\{A\}+P\{B_1\}+P\{B_2\}+P\{B_3\}+P\{B_4\}=0+\frac{1}{m(m-1)}+\frac{m-3}{m(m-1)(m-2)}+0+\frac{(m-4)(m-3)}{m(m-1)(m-2)}$$

$$=\frac{(m-3)^2}{m(m-1)(m-2)}+\frac{1}{m(m-1)}=\frac{1}{m}-\frac{2}{m(m-1)}+\frac{1}{m(m-1)(m-2)}.$$

The lemma is proved.

Lemma 14. *Let $i_1=2j_0$. Then*

a) $P\{\, z_1=v\,\}=\frac{1}{m}-\frac{1}{m(m-1)}$ *for $v\neq j_0$,*

b) $P\{z_1=j_0\}=\frac{2}{m}$.

Proof. Let us consider the following cases:

1. if $v\neq j_0$, then

$$P\{z_1=v\}=\frac{1}{m(m-1)}+0+0+\frac{m-3}{m(m-1)(m-2)}+\frac{(m-3)^2}{m(m-1)(m-2)}=\frac{(m-2)}{m(m-1)}=\frac{1}{m}-\frac{1}{m(m-1)}.$$

2. if $v=j_0$, then

$$P\{z_1=j_0\}=\frac{1}{m}+0+0+0+\frac{1}{m}=\frac{2}{m}.$$

The lemma is proved.

Lemmas 10-14 complete the proof.

Theorem 3 describes the distribution of the second output value z_2.

Theorem 3. **(the distribution of the second output value z_2)** *Assume that the permutation $s_0\in S_m$ is randomly chosen from S_m. Let $(i_0, j_0, s_0)\in Z_m\times Z_m\times S_m$ be any initial state of RC4. Then:*

I.

a) $P\{z_2=0\}=\frac{3}{m}+O(\frac{1}{m^2})$ *for $i_2=j_0=0$,*

b) $P\{z_2=k\}=\frac{1}{m}+O(\frac{1}{m^2})$ *for $i_2=j_0=0$, $k\neq 0$.*

II.

a) $P\{z_2=0\}=\frac{2}{m}+O(\frac{1}{m^2})$ *for $j_0=0$, $i_2\neq 0$,*

b) $P\{z_2=k\}=\frac{1}{m}+O(\frac{1}{m^2})$ *for $j_0=0$, $i_2\neq 0$, $k\neq 0$.*

Another cases $P\{z_2=k\}=\frac{1}{m}+O(\frac{1}{m^2})$, $k=\overline{0,m-1}$.

The theorem is proved as theorem 2. Stress that the first special case ($P\{z_2=0\}=\frac{2}{m}$, where $j_0=i_0=0$, $i_2=2$) of theorem 3 was found A. Shamir, I. Mantin [3].

5. The distribution of digraphs of RC4

In this section we find the distribution of digraphs in an output sequence of RC4. We begin with the following theorem.

Theorem 4 (conditional probabilities of the second output value) *Assume that the permutation $s_0 \in S_m$ is randomly chosen from S_m. Let $(i_0, j_0, s_0) \in Z_m \times Z_m \times S_m$ be an initial state of RC4.*

I. Let $j_0=0$, $i_2=0$. Then

a) $P\{z_2=0 \,|z_1=k_1\}=\frac{3}{m}+O(\frac{1}{m^2})$ for $k_1 \neq 0$,

b) $P\{z_2=k_2 \mid z_1=k_1\}=\frac{1}{m}+O(\frac{1}{m^2})$ for $k_2 \neq 0$.

II. Let $j_0=0$, $i_2 \neq 0$. Then

a) $P\{z_2=0 \,|z_1=k_1\}=\frac{2}{m}+O(\frac{1}{m^2})$ for $k_1 \neq 0$,

b) $P\{z_2=k_2 \mid z_1=k_1\}=\frac{1}{m}+O(\frac{1}{m^2})$ for $k_2 \neq 0$.

III. Let $j_0 \neq 0$. Then

a) $P\{z_2=k_2 \mid z_1=0\}=\frac{2}{m}+O(\frac{1}{m^2})$ for $k_1=0$, $k_2=i_1=j_0$, $k_2 \neq i_2$, $i_2=2j_0 (mod\ m)$,

b) $P\{z_2=k_2 \mid z_1=k_1\}=\frac{1}{m}+O(\frac{1}{m^2})$ for another cases .

Proof. Note that $k_1=s_1[s_1[i_1]+s_1[j_1]]=s_1[j_2-j_0-s_1[i_2]+s_1[i_1]]$ and consider random variable: $j_2=s_1[i_2]+s_1[j_1]+j_0$ и $\gamma_2=s_1[i_2]+s_1[j_2]$. It is clear that they are dependent.

From the full probability formula we get that

$$\begin{aligned}
P\{z_2=k_2|\, z_1=k_1\}= &\sum_{\gamma=0}^{m-1} P\{\gamma_2=\gamma\}P\{s_2[\gamma]=k_2 \mid s_1[\gamma-j_0-s_1[i_2]+s_1[i_1]]=k_1\}=\sum_{\substack{\gamma=0\\ \gamma\neq i_2}}^{m-1}\ \sum_{\substack{r=0\\ r\neq\gamma,i_2}}^{m-1} P\{\gamma_2=\gamma\}P\{j_2=r \mid \\
&\gamma_2=\gamma\}\, P\{s_1[\gamma]=k_2|\, s_1[\gamma-j_0-s_1[i_2]+s_1[i_1]]=k_1\}+\sum_{\substack{\gamma=0\\ \gamma\neq i_2}}^{m-1} P\{\gamma_2=\gamma\}\, P\{j_2=\gamma \mid \gamma_2=\gamma\}P\{s_1[i_2]=k_2|\, s_1[\gamma-j_0-s_1[i_2] \\
&+s_1[i_1]]=k_1\}+\sum_{\substack{\gamma=0\\ \gamma\neq i_2}}^{m-1} P\{\gamma_2=\gamma\}P\{j_2=i_2 \mid \gamma_2=\gamma\}P\{s_1[\gamma]=k_2 \mid s_1[\gamma-j_0-s_1[i_2]+s_1[i_1]]=k_1\}+\sum_{\substack{r=0\\ r\neq i_2}}^{m-1} P\{\gamma_2=i_2\}\, P\{j_2=r \\
&\mid \gamma_2=i_2\}\, P\{s_1[r]=k_2|\, s_1[i_2-j_0-s_1[i_2]+s_1[i_1]]=k_1\}+P\{\gamma_2=i_2\}P\{j_2=i_2|\gamma_2=i_2\}P\{s_1[i_2]=k_2|\, s_1[i_2-j_0-s_1[i_2]+ \\
&s_1[i_1]]=k_1\}. \qquad (4)
\end{aligned}$$

From (4), it follows that $P\{z_2=k_2|\, z_1=k_1\}=P\{A_1\}+P\{A_2\}+P\{A_3\}+P\{A_4\}+P\{A_5\}$, where

$$P\{A_1\}=\sum_{\substack{\gamma=0\\ \gamma\neq i_2}}^{m-1}\ \sum_{\substack{r=0\\ r\neq\gamma,i_2}}^{m-1} P\{\gamma_2=\gamma\}P\{j_2=r \mid \gamma_2=\gamma\}\, P\{s_1[\gamma]=k_2|\, z_1=k_1\},$$

$$P\{A_2\}=\sum_{\substack{\gamma=0\\ \gamma\neq i_2}}^{m-1} P\{\gamma_2=\gamma\}\, P\{j_2=\gamma \mid \gamma_2=\gamma\}P\{s_1[i_2]=k_2|\, z_1=k_1\},$$

$$P\{A_3\}=\sum_{\substack{\gamma=0\\ \gamma\neq i_2}}^{m-1} P\{\gamma_2=\gamma\}P\{j_2=i_2 \mid \gamma_2=\gamma\}P\{s_1[\gamma]=k_2 \mid z_1=k_1\},$$

$$P\{A_4\}=\sum_{\substack{r=0\\ r\neq i_2}}^{m-1} P\{\gamma_2=i_2\}P\{j_2=r \mid \gamma_2=i_2\}\ P\{s_1[r]=k_2 \mid z_1=k_1\},$$

$$P\{A_5\}=P\{\gamma_2=i_2\}P\{j_2=i_2 \mid \gamma_2=i_2\}P\{s_1[i_2]=k_2 \mid z_1=k_1\}.$$

Denote $s_1[i_1]=h$. We shall estimate of $P\{A_1\},\ldots, P\{A_5\}$.

Lemma 1 $P\{A_1\}=\frac{1}{m}+O(\frac{1}{m^2})$.

Proof. Note that

$$P\{A_1\}=\sum_{\substack{\gamma=0\\ \gamma\neq i_2}}^{m-1}\ \sum_{\substack{r=0\\ r\neq\gamma,i_2}}^{m-1} P\{\gamma_2=\gamma\}P\{j_2=r \mid \gamma_2=\gamma\}\ P\{s_1[\gamma]=k_2 \mid z_1=k_1\}=\sum_{h=0}^{m-1}\ \sum_{\substack{\gamma=0\\ \gamma\neq i_2}}^{m-1}\ \sum_{\substack{r=0\\ r\neq\gamma,i_2}}^{m-1}\ \sum_{\substack{t=0,\\ t\neq k_2,\\ \gamma\neq 2t}}^{m-1} P\{s_1[i_1]=h \mid$$

$s_1[\gamma-j_0-t+h]=k_1\}$ $P\{s_1[\gamma]=k_2 \mid s_1[\gamma-j_0-t+h]=k_1, s_1[i_1]=h\}$ $P\{s_1[i_2]=t \mid s_1[\gamma-j_0-t+h]=k_1\}$ $P\{s_1[r]=\gamma-t \mid s_1[i_2]=t, s_1[\gamma]=k_2, s_1[\gamma-j_0-t+h]=k_1, s_1[i_1]=h\}P\{s_1[r-t]=r-t-j_0 \mid s_1[r]=\gamma-t, s_1[i_2]=t, s_1[\gamma]=k_2, s_1[\gamma-j_0-t+h]=k_1, s_1[i_1]=h\}$.

Therefore,

$$P\{A_1\}\sim\frac{m(m-1)(m-2)(m-3)}{(m-1)(m-2)(m-3)(m-4)(m-5)}=\frac{1}{m}+O(\frac{1}{m^2});$$

The lemma is proved.

Lemma 2. *If $j_0=0$, $k_1\neq 0$, $k_2=0$, then*

$$P\{A_2\}=\frac{1}{m}+O(\frac{1}{m^2}).$$

else

$$P\{A_2\}=O(\frac{1}{m^2}).$$

Proof. Note that

$$P\{A_2\}=\sum_{\substack{\gamma=0\\ \gamma\neq i_2}}^{m-1} P\{\gamma_2=\gamma\}\ P\{j_2=\gamma \mid \gamma_2=\gamma\}P\{s_1[i_2]=k_2 \mid z_1=k_1\}=\sum_{\substack{\gamma=0\\ \gamma\neq i_2}}^{m-1} P\{s_1[i_2]=k_2 \mid z_1=k_1\}\sum_{t=0}^{m-1} P\{s_1[i_2]=t \mid$$

$s_1[i_2]=k_2, z_1=k_1\}$ $P\{s_1[\gamma]=\gamma-t \mid s_1[i_2]=t, s_1[i_2]=k_2, z_1=k_1\}P\{s_1[\gamma-t]=\gamma-t-j_0 \mid s_1[\gamma]=\gamma-t, s_1[i_2]=t,$

$$s_1[i_2]=k_2,\ z_1=k_1\}=\sum_{\substack{h=0\\ h\neq k_2}}^{m-1}\ \sum_{\substack{\gamma=0\\ \gamma\neq i_2,2k_2}}^{m-1} P\{s_1[i_2]=k_2 \mid s_1[i_2-j_0-k_2+h]=k_1\}\ P\{s_1[i_1]=h \mid s_1[i_2-j_0-k_2+h]=k_1\}$$

$P\{s_1[\gamma]=\gamma-k_2 \mid s_1[i_2]=k_2, s_1[i_2-j_0-k_2+h]=k_1, s_1[i_1]=h\}P\{s_1[\gamma-k_2]=\gamma-k_2-j_0 \mid s_1[\gamma]=\gamma-k_2, s_1[i_2]=k_2, s_1[i_2-j_0-k_2+h]=k_1, s_1[i_1]=h\}$.

Let us consider the following cases.

I. Let $j_0=0$; then:

a) for $k_1\neq 0$, $k_2=0$ we get

$$P\{A_2\}=\sum_{\substack{h=0\\ h\neq 0}}^{m-1}\ \sum_{\substack{\gamma=0\\ \gamma\neq i_2,0}}^{m-1} P\{s_1[i_2]=0 \mid s_1[i_2+h]=k_1\}\ P\{s_1[i_1]=h \mid s_1[i_2+h]=k_1\}\ P\{s_1[\gamma]=\gamma \mid s_1[i_2]=0,$$

$$s_1[i_2+h]=k_1,\ s_1[i_1]=h\}=\sum_{\substack{h=0\\ h\neq 0}}^{m-1}\ \sum_{\substack{\gamma=0,\gamma\neq i_2+h\\ \gamma\neq i_2,0,k_1}}^{m-1} P\{s_1[i_2]=0 \mid s_1[i_2+h]=k_1\}\ P\{s_1[i_1]=h \mid s_1[i_2+h]=k_1\}\ P\{s_1[\gamma]=\gamma \mid$$

$s_1[i_2]=0$, $s_1[i_2+h]=k_1$, $s_1[i_1]=h\}+$ $P\{s_1[i_2]=0 \mid s_1[k_1]=k_1\}$ $P\{s_1[i_1]=k_1-i_2 \mid s_1[i_2+h]=k_1\}\sim$

$\frac{1}{(m-1)(m-2)}+\frac{(m-1)(m-4)}{(m-1)(m-2)(m-3)}\sim\frac{1}{(m-1)}+\frac{(m-1)(m-4)}{(m-1)(m-2)(m-3)}\sim\frac{1}{(m-1)}+\frac{1}{(m-1)(m-2)}=$ $\frac{1}{m}+\mathrm{O}(\frac{1}{m^2})$.

c) for $k_1= k_2=0$ we get

$P\{A_2\}=\sum_{\substack{h=0\\h\neq 0}}^{m-1}\ \sum_{\substack{\gamma=0\\ \gamma\neq i_2,0}}^{m-1} P\{s_1[i_2]=0|\ s_1[i_2+h]=0\}\ P\{s_1[i_1]=h\ |\ s_1[i_2+h]=0\}\ P\{s_1[\gamma]=\gamma\ |\ s_1[i_2]=0, s_1[i_2+h]=k_1,$ $s_1[i_1]=h\}=0$;

b) for $k_2\neq 0$ we have

$P\{A_2\}=\sum_{\substack{h=0\\h\neq k_2}}^{m-1}\ \sum_{\substack{\gamma=0\\ \gamma\neq i_2,2k_2}}^{m-1} P\{s_1[i_2]=k_2|\ s_1[i_2-k_2+h]=k_1\}\ P\{s_1[i_1]=h\ |\ s_1[i_2-k_2+h]=k_1\}\ P\{s_1[\gamma]=\gamma-k_2\ |\ s_1[i_2]$ $=k_2, s_1[i_2-k_2+h]=k_1, s_1[i_1]=h\}P\{s_1[\gamma-k_2]=\gamma-k_2|\ s_1[\gamma]=\gamma-k_2, s_1[i_2]=k_2, s_1[i_2-k_2+h]=k_1, s_1[i_1]=h\}=0$.

II. Let $j_0\neq 0$; then

$P\{A_2\}=\sum_{\substack{h=0\\h\neq k_2}}^{m-1}\ \sum_{\substack{\gamma=0\\ \gamma\neq i_2,2k_2}}^{m-1} P\{s_1[i_2]=k_2|\ s_1[i_2-j_0-k_2+h]=k_1\}\ P\{s_1[i_1]=h\ |\ s_1[i_2-j_0-k_2+h]=k_1\}\ P\{s_1[\gamma]=\gamma-k_2\ |$ $s_1[i_2]=k_2, s_1[i_2-j_0-k_2+h]=k_1, s_1[i_1]=h\}P\{s_1[\gamma-k_2]=\gamma-k_2-j_0\ |\ s_1[\gamma]=\gamma-k_2, s_1[i_2]=k_2, s_1[i_2-j_0-k_2+h]=$ $k_1, s_1[i_1]=h\}$.

a) for $k_2=0$ we get $P\{A_2\}=0$.

b) for $k_2\neq 0$ we have $P\{A_2\}\sim\frac{1}{(m-1)(m-2)}=\mathrm{O}(\frac{1}{m^2})$.

The lemma is proved.

Lemma 3. *If $i_2=j_0=k_1= k_2=0$, then $P\{A_5\}=\frac{1}{(m-1)}$, else $P\{A_5\}=O(\frac{1}{m^2})$.*

Proof. We stress that

$P\{A_5\}=\sum_{h=0}^{m-1} P\{\gamma_2=i_2\}P\{j_2=i_2\ |\ \gamma_2=i_2\}P\{\ s_1[i_1]=h\ |\ s_1[i_2-j_0-s_1[i_2]+h]=k_1\}P\{s_1[i_2]=\ k_2|\ s_1[i_2-j_0-s_1[i_2]$ $+h]=k_1\}=\sum_{h=0}^{m-1} P\{s_1[i_2]=k_2|\ s_1[i_2-j_0-s_1[i_2]+h]=k_1\}\ P\{\ s_1[i_1]=h\ |\ s_1[i_2-j_0-s_1[i_2]+h]=k_1,\ s_1[i_2]=k_2\}$ $P\{s_1[i_2]=i_2-s_1[i_2]\ |\ s_1[i_2]=k_2,\ s_1[i_2-j_0-s_1[i_2]+h]=\ k_1,\ s_1[i_1]=h\}P\{s_1[i_2-\ s_1[i_2]]=i_2-s_1[i_2]-j_0\ |\ s_1[i_2]=k_2,$ $2k_2=i_2, s_1[i_2-j_0-s_1[i_2]+h]=k_1, s_1[i_2-j_0-s_1[i_2]+h]=k_1, s_1[i_1]=h\}= P\{s_1[i_2]=k_2|\ s_1[i_1]=k_1, s_1[i_2]=i_2-j_0+ k_1\}$ $P\{s_1[k_2]=k_2-j_0\ |\ s_1[i_2]=k_2,\ 2k_2=i_2,\ s_1[i_1]=\ k_1\}+\sum_{\substack{h=0\\h\neq k_2,k_1}}^{m-1} P\{\ s_1[i_1]=h\ |\ s_1[i_2-j_0-s_1[i_2]+h]=k_1\}\ P\{s_1[i_2]=k_2|$ $s_1[i_1]=h, s_1[i_2-j_0-s_1[i_2]+h]=k_1\}\ P\{s_1[k_2]=k_2-j_0\ |\ s_1[i_1]=h, s_1[i_2]=k_2, 2k_2=i_2, s_1[i_2-j_0-s_1[i_2]+h]=k_1\}$.

Therefore, $P\{A_5\}= P\{B_1\}+ P\{B_2\}$, where

$P\{B_1\}= P\{s_1[i_2]=k_2|\ s_1[i_1]=k_1, s_1[i_2]=i_2-j_0+ k_1\}\ P\{s_1[k_2]=k_2-j_0\ |\ s_1[i_2]=k_2, 2k_2=i_2, s_1[i_1]= k_1\}$,

$P\{B_2\}=\sum_{\substack{h=0\\h\neq k_2,k_1}}^{m-1} P\{\ s_1[i_1]=h\ |\ s_1[i_2-j_0-s_1[i_2]+h]=k_1\}\ P\{s_1[i_2]=k_2|\ s_1[i_1]=h, s_1[i_2-j_0-s_1[i_2]+h]=k_1\}P\{s_1[k_2]$ $=k_2-j_0\ |\ s_1[i_1]=h, s_1[i_2]=k_2, 2k_2=i_2, s_1[i_2-j_0-s_1[i_2]+h]=k_1\}$.

We will consider the following cases.

I. Let $2k\neq i_2$; then $P\{A_5\}=0$.

II. Let $j_0=0$; then:

a) for $k_2=i_2$; $2k_2=i_2$; $k_2=i_2=0$, $k_1\neq 0$ we have

$$P\{B_1\}= P\{s_1[i_2]=0|\ s_1[i_1]=k_1,\ s_1[i_2]=k_1\}\ P\{s_1[k_2]=k_2\ |\ s_1[i_2]=k_2,\ 2k_2=i_2,\ s_1[i_1]= k_1\}=0.$$

$$P\{B_2\}= \sum_{\substack{h=0\\ h\neq 0,i_1-i_2}}^{m-1} P\{s_1[i_1]=h\ |\ s_1[i_2+h]=k_1\}\ P\{s_1[i_2]=0|\ s_1[i_1]=h,\ s_1[i_2+h]=k_1\}=\frac{1}{(m-1)}.$$

Therefore, $P\{A_5\}=\frac{1}{(m-1)}$.

b) for $k_2\neq i_2$ we get that $P\{A_5\}=0$.

II. Let $j_0\neq 0$; then:

a) for $k_2\neq i_2-j_0+ k_1$ we have

$$P\{B_1\}=P\{s_1[i_2]=k_2|\ s_1[i_1]=k_1,\ s_1[i_2]=i_2-j_0+ k_1\}\ P\{s_1[k_2]=k_2-j_0\ |\ s_1[i_2]=k_2,\ 2k_2=i_2,\ s_1[i_1]= k_1\}=0.$$

b) for $k_2= i_2-j_0+ k_1$; $k_2\neq i_2$; $i_1=k_2$; $2k_2=i_2$; $i_2=2j_0$; $i_2=2(j_0- k_1)$; $k_1=0$; $k_2=j_0$ we get

$$P\{B_1\}=P\{s_1[i_2]= i_2-j_0+ k_1|\ s_1[i_1]=k_1\}=\frac{1}{(m-1)}.$$

c) for $k_2= i_2-j_0+ k_1$; $k_2\neq i_2$; $i_1\neq k_2$; $2k_2=i_2$; i.e. $i_2=2(j_0- k_1)$; $k_1\neq 0$ we have

$$P\{B_1\}=P\{s_1[i_2]=i_2-j_0+k_1|\ \ s_1[i_1]=k_1\}\ \ P\{s_1[k_2]=i_2-2j_0+k_1\ \ |\ \ s_1[i_2]=i_2-j_0+k_1,\ \ 2k_2=i_2,\ \ s_1[i_1]=k_1\}= \frac{1}{(m-1)(m-2)}=O(\frac{1}{m^2}).$$

In other cases $P\{B_1\}=0$.

III. Let $k_1=k_2$; $j_0\neq 0$; $2k_2=i_2$; ($s_1[i_1]=h=k_2+j_0$); then:

a) for $k_2= i_1$ we get

$$P\{B_2\}=P\{s_1[i_1]= k_2+j_0\ |\ s_1[i_2]=k_2\}\ P\{s_1[k_2]=k_2-j_0\ |\ s_1[i_2]=k_2,\ 2k_2=i_2,\ s_1[i_1]= k_2+j_0\}=0.$$

b) for $k_2\neq i_1$ we have

$$P\{B_2\}=P\{s_1[i_1]= k_2+j_0\ |\ s_1[i_2]=k_2\}\ P\{s_1[k_2]=k_2-j_0\ |\ s_1[i_2]=k_2,\ 2k_2=i_2,\ s_1[i_1]= k_2+j_0\}= \frac{1}{(m-1)(m-2)}= O(\frac{1}{m^2}).$$

IV. Let $k_1\neq k_2$; $2k_2=i_2$; then:

a) for $k_2=i_2=0$ we have

$$P\{B_2\}= \sum_{\substack{h=0\\ h\neq k_2,k_1}}^{m-1} P\{\ \ s_1[i_1]=h\ \ |\ \ s_1[i_2-j_0-s_1[i_2]+h]=k_1\}\ \ P\{s_1[i_2]=k_2|\ \ s_1[i_1]=h,\ \ s_1[i_2-j_0-s_1[i_2]+h]=k_1\}$$

$$P\{s_1[k_2]=k_2-j_0\ |\ s_1[i_1]=h,\ s_1[i_2]=k_2,\ 2k_2=i_2,\ s_1[i_2-j_0-s_1[i_2]+h]=k_1\}=0.$$

b) for $k_2=i_1$ we have

$$P\{B_2\}=P\{s_1[i_2]=k_2|\ s_1[i_1]= k_2-j_0,\ s_1[i_2-j_0-s_1[i_2]+h]=k_1\}\ P\{s_1[k_2]=k_2-j_0\ |\ s_1[i_2]=k_1\}=0.$$

c) for $k_2\neq 0$, $i_2\neq 0$; $j_0\neq 0$ we get

$$P\{B_2\}= \sum_{\substack{h=0\\ h\neq k_2,k_1}}^{m-1} P\{\ \ s_1[i_1]=h\ \ |\ \ s_1[i_2-j_0-s_1[i_2]+h]=k_1\}\ \ P\{s_1[i_2]=k_2|\ \ s_1[i_1]=h,\ \ s_1[i_2-j_0-s_1[i_2]+h]=k_1\}$$

$$P\{s_1[k_2]= k_2-j_0\ |\ s_1[i_1]=h,\ s_1[i_2]=k_2,\ 2k_2=i_2,\ s_1[i_2-j_0-s_1[i_2]+h]=k_1\}\sim\frac{1}{(m-1)(m-2)}= O(\frac{1}{m^2}).$$

The lemma is proved.

Lemma 4.

a) $P\{A_3\}=O(\frac{1}{m^2})$,

b) $P\{A_4\}= O(\frac{1}{m^2})$.

The proof is by direct calculation.

The proof follows from lemmas 1–4 and $P\{z_2=k_2|\ z_1=k_1\}=P\{A_1\}+P\{A_2\}+P\{A_3\}+P\{A_4\}+P\{A_5\}$. The theorem is proved.

Note that events A_2 and A_5 give non-uniformity in the distribution.

$$P\{A_2\}=\sum_{\substack{\gamma=0\\ \gamma\neq i_2}}^{m-1} P\{\gamma_2=\gamma\}\ P\{j_2=\gamma\mid \gamma_2=\gamma\}P\{s_1[i_2]=k_2|\ z_1=k_1\};$$

$$P\{A_5\}=P\{\gamma_2=i_2\}P\{j_2=i_2\mid \gamma_2=i_2\}P\{s_1[i_2]=k_2|\ z_1=k_1\}.$$

1. The event A_2 means that we have $s_2[i_2]=s_1[j_2]=k_2$, $s_1[i_2]+s_1[j_2]=s_2[i_2]+s_2[j_2]=j_2$ and $s_1[i_2]+s_1[j_2]\neq i_2$.
2. The event A_5 means that we have $j_2=i_2=s_1[i_2]+s_1[j_2]=s_2[i_2]+s_2[j_2]$ and $s_1[i_2]=s_1[j_2]=s_2[i_2]=s_2[j_2]$.

By the previous theorem and theorem 2 we obtain the proof theorem 5. We stress that the following result was presented at the MEPhI's [10].

Theorem 5. (the distribution of digraphs)

Assume that the permutation $s_0\in S_m$ is randomly chosen from S_m. Let $(i_0, j_0, s_0)\in Z_m\times Z_m\times S_m$ be any initial state of RC4.

I. Let $i_1=j_0=0$. Then:

a) $P\{z_1=0, z_2=k_2\}=\frac{2}{m^2}+O(\frac{1}{m^3})$ *for* $k_1=0, k_2\neq 0$,

b) $P\{z_1=k_1, z_2=0\}=\frac{2}{m^2}+O(\frac{1}{m^3})$ *for* $k_1\neq 0$,

c) $P\{z_1=k_1, z_2=k_2\}=\frac{1}{m^2}+O(\frac{1}{m^3})$ *for* $k_1\neq 0, k_2\neq 0$.

II. Let $j_0=0$, $i_1\neq 0$. Then:

a) $P\{z_1=k_1, z_2=0\}=\frac{2}{m^2}+O(\frac{1}{m^3})$ *for* $k_2=0, k_1\neq 0$,

b) $P\{z_1=k_1, z_2=k_2\}=\frac{1}{m^2}+O(\frac{1}{m^3})$ *for* $k_2\neq 0$.

III. Let $i_1=2j_0$, $j_0\neq 0$, $i_1\neq j_0$. Then:

a) $P\{z_1=j_0, z_2=k_2\}=\frac{2}{m^2}+O(\frac{1}{m^3})$ *for* $k_1=j_0, k_2\neq 0$,

b) $P\{z_1=k_1, z_2=k_2\}=\frac{1}{m^2}+O(\frac{1}{m^3})$ *for* $k_1\neq j_0, k_2\neq 0$.

IV. Let $i_2=2j_0$, $i_1=j_0$, $j_0\neq 0$. Then:

a) $P\{z_1=0, z_2=j_0\}=\frac{2}{m^2}+O(\frac{1}{m^3})$ *for* $k_1=0, k_2=j_0$,

b) $P\{z_1=k_1, z_2=k_2\}=\frac{1}{m^2}+O(\frac{1}{m^3})$ *for* $k_1\neq 0, k_2\neq j_0$.

V. Let $i_2=j_0=0$, $i_1\neq i_2$. Then:

1. $P\{z_1=k_1, z_2=0\}=\frac{3}{m^2}+O(\frac{1}{m^3})$ *for* $k_2=0, k_1\neq 0$,
2. $P\{z_1=k_1, z_2=k_2\}=\frac{1}{m^2}+O(\frac{1}{m^3})$ *for* $k_2\neq 0, k_1\neq 0$.

Another cases $P\{z_1=k_1,z_2=k_2\}=\frac{1}{m^2}+O(\frac{1}{m^3})$, $(k_1, k_2)\in Z_m^2$.

6. Conclusion

In this paper we considered statistical properties of the RC4 stream cipher. We proved that the distribution of first, second output values of RC4 and digraphs are not uniform. This makes RC4 trivial to distinguish between short outputs of RC4 and random strings by analyzing their first, or second output values of RC4 or digraphs.

7. Acknowledgments

The author is grateful to my scientific adviser professor Pogorelov B. A. for constant attention to this work.

References.

[1] Golic, J. D, Linear Statistical Weakness of Alleged RC4 Keystream Generator. Advances in Cryptology -- EUROCRYPT '97.

[2] Fluhrer S.R., McGrew D. A. “Statistical analysis of the alleged RC4 keystream generator”, Proceeding of FSE’2000, Springer-Verlag.

[3] Mantin I. Shamir A. “A practical attack on broadcast RC4”, Proceeding of FSE’2001, Springer-Verlag.

[4] Mister S., Tavares S., “Cryptanalysis of RC4-like ciphers”, Proceeding of SAC’98, Springer-Verlag.

[5] Knudsen L., Meier W., Preneel B., Rijmen V., Verdoolaege S, “Analysis method for (alleged) RC4”, Proceeding of ASIACRYPT’99, Springer-Verlag.

[6] Grosul A.L., Wallach D.S. “A related key cryptanalysis of RC4”, 2000, to appear.

[7] Pudovkina M. “Short cycles of the alleged RC4 keystream generator ”, 3nd International Workshop on Computer Science and Information Technologies, CSIT’2001, YFA, 2001.

[8] M. Pudovkina “Statistical weakness in the alleged RC4 keystream generator”, 4 International Workshop on Computer Science and Information Technologies, CSIT’2002, 2002.

[9] Пудовкина М. А. “О распределении первого выходного символа криптосистеме RC4” , в сб. научных трудов XLIV юбилейной научной конференции МФТИ, Москва – Долгопрудный, 2001.(in Russian)

[10] Пудовкина М. А “О распределении биграмм в криптосхеме RC4”, В сб. научных трудов конференции «Проблемы информационной безопасности в системе высшей школы», Москва, 2002. (in Russian)

[11] Пудовкина М. А. “Об одной системе образующих с ограничениями”, В сб. научных трудов конференции «Проблемы информационной безопасности в системе высшей школы», Москва, 2002. (in Russian)

[12] Пудовкина М. А. “О свойствах алгоритма поточного шифрования RC4”. В сб. тезисов конференции «Методы и технические средства обеспечения безопасности информации», Санкт-Петербург, 2001. (in Russian)